# cNORM Weighting Manuscript: Editorial Discussion

Before I start editing, I want to make sure I understand the method of the study, and also address nomenclature for the various aspects of the method.

# Descriptive Labels

One thing that would make the manuscript more readable is to replace the current *non-descriptive labeling* of various aspects of the study.

* Stratification variables: variable names are 1,2,3, and variable levels are 1,2,3
* Sampling conditions are labeled 1-6
* The 27 strata have 1,1,1 (etc.) subscripts
* In the post-hoc analysis, the ANOVAs are referred to as the first and second ANOVAs

The use of non-descriptive labels (integers) for all conditions and levels can make it hard to follow the narrative as you describe this complex methodology.

Here are some ideas for descriptive labels:

## Stratification Variables:

* Line 698 refers to the SVs in this study being modeled after *education*, *ethnicity*, *region* – what if we used these as labels?
  + Levels could then be:
    - *no high school diploma, high school grad, college degree*
    - *White, Black, Hispanic*
    - *South, North-East, West*
    - (or something similar)

## Sampling Conditions:

1. Unbiased
2. Small bias low ability
3. Large bias low ability
4. Bias low and high ability
5. Bias joint probabilities
6. Age-related bias

# Clarification of Methods:

## Reference Population

The simulated reference population of 24 million persons is first described in detail starting on line 342. Then starting on line 360 you refer to the creation of five additional population samples “in the same way”.

* Does this mean that the five “population samples” (description starts on line 366) are also 24 million persons, with 4 million per age cohort?
* If this is the case, then I suggest we refer to all six “population samples” as “simulated populations”. This maintains the conventional distinction between a “population” (the entire collection of people) from which we draw “samples” (which may be more or less representative of the “population”).
* If what I've said so far is correct, then the clearest description of the data hierarchy is:
  + Six simulated populations (n = 24 million).
  + 100 normative samples (n = 600) are drawn from each simulated population.

## Stratification Variables

**Table 1**

*Proportions of the levels of the three stratification variables in the reference population*

|  |  |  |  |
| --- | --- | --- | --- |
| Stratification variable | Level 1 | Level 2 | Level 3 |
| 1 | 40 % | 20 % | 40 % |
| 2 | 30 % | 40 % | 30 % |
| 3 | 60 % | 20 % | 20 % |

Table 1 anchors much of the method description, but it is hard to decipher.

* Each SV has three levels
  + What does “level” mean in this context? I'm not clear as to whether it means something like a demographic category (e.g., black vs. white ethnicity) or what the manuscript refers to as “performance” (1 = above, 2 = average, 3 = below)?
* The next paragraph refers to “3 x 3 x 3 = 27 strata”, but this description is at odds with the structure of three SVs, each with three levels, which would lead to 3 x 3 = 9 strata or joint categories.
* Part of the reason why Table 1 is confusing is that it combines information on both *structure* (in the table margins) and *distribution* (in the cells).
* The 3 x 3 x 3 structure can be shown more clearly in a nested table that shows only classification structure:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | SV1 | SV2 | SV3 |
| SV1 | SV1 | 1,1,1 | 2,1,1 | 3,1,1 |
| SV2 | 1,1,2 | 2,1,2 | 3,1,2 |
| SV3 | 1,1,3 | 2,1,3 | 3,1,3 |
| SV2 | SV1 | 1,2,1 | 2,2,1 | 3,2,1 |
| SV2 | 1,2,2 | 2,2,2 | 3,2,2 |
| SV3 | 1,2,3 | 2,2,3 | 3,2,3 |
| SV3 | SV1 | 1,3,1 | 2,3,1 | 3,3,1 |
| SV2 | 1,3,2 | 2,3,2 | 3,3,2 |
| SV3 | 1,3,3 | 2,3,3 | 3,3,3 |

* Here, the cell values (1,1,1 . . . etc.) represent the 27 possible combinations of the levels of the three SVs. Of course “collegeGrad-black-west” would be a lot more meaningful and vivid than “3,2,3” (for example).
* Additional implicit three-category classifications related to the SVs:
  + Effect of SV on the latent cognitive ability (*large*, *medium*, *small*). As I understand it, this correlation between SV and ability differs between SVs, but is constant within the levels of each SV?
  + *Above* vs. *average* vs. *below* performance. This presumably refers to the latent cognitive ability, in the form of the group mean of the raw test score from the simulated test?
  + Distribution of levels of SVs within the reference population (*40-20-40* vs. *30-40-30* vs. *60-20-20*, as shown in Table 1).
* With all of this information flying around, it's not clear to me what values belong in the cells of my 27-cell nested table. To me, this is the most important issue to clarify in the methods section of this paper.
* Another way to configure the 27-cell table is to relate it back to the descriptions of the sampling methods.
  + The primary thing that's being manipulated to differentiate the six sampling conditions is the distribution of the levels of SV1 (large effect on cognitive ability).
  + Thus, if you applied my 27-cell table to each of the six sampling conditions, what numbers should we put in the 27 cells to most clearly show the difference between the sampling conditions?

### Interaction between weighting and person location: First and Second ANOVAs

Here's what I can figure out (but need confirmation) about this aspect of the design.

ANOVA configuration:

* First: 11 (location) X 2 (norming method)
* Second: 11 (location) x 6 (sampling condition)

ANOVA outcome:

* First: Compare WCN to SCN across the range of person abilities: which method yielded norm scores that were closer to the benchmark (unbiased population sample)?
* Second: How did the accuracy of WCN (with respect to the benchmark) change across the range of person abilities?

Do I have all this correct?